



# **Analyse av lønnsforskjeller mellom kvinner og menn i Italia og Norge**

Bacheloroppgave i samfunnsøkonomi

SØK2901

NORGES TEKNISK- NATURVITENSKAPELIGE UNIVERSITET

Trondheim, Våren 2019

## **Sammendrag**

Innen forskning har interessen for yrkesaktive kvinner økt nasjonalt og internasjonalt de senere årene. Kvinner har stadig blitt en større del av arbeidsstyrken, og følgelig har lik lønn mellom kvinner og menn blitt en viktig politisk målsetting. I denne bacheloroppgaven har vi undersøkt om det eksisterer et lønnsgap mellom kvinner og menn i to europeiske land: Italia og Norge. Vi har sett på hva som påvirker kvinner og menns lønnsnivå ved å benytte data fra den internasjonale undersøkelsen, PIACC, som er gjennomført av Organisasjonen for økonomisk samarbeid og utvikling (OECD).

Vi har undersøkt hvordan både humankapitalvariabler og arbeidsrelaterte variabler påvirker lønnsnivået til kvinner og menn. I oppgaven observerer vi et lønnsgap mellom kvinner og menn i Italia og Norge samt lønnsforskjeller mellom de to landene. Videre observerer vi et lønnsgap mellom privat og offentlig sektor, og finner at timelønnen er høyest for norske menn som jobber i privat sektor. I Italia er derimot timelønnen høyest for menn ansatt i offentlig sektor. Marginalavkastningen av utdanning og kognitive ferdigheter er lik for kjønnene i begge land, lønnsgapet forklares av blant annet lavere forventet startlønn for kvinner.

## Innholdsfortegnelse

1	Innledning.....	4
1.1	Problemstilling .....	4
1.2	Struktur.....	5
2	Teoretisk rammeverk og tidligere litteratur.....	6
3	Datamaterialet .....	8
3.1	Om datamaterialet .....	8
3.2	Deskriptive analyser .....	9
3.2.1	Deskriptiv statistikk for avhengig variabel .....	10
3.2.2	Deskriptiv statistikk for uavhengige variabler .....	11
3.3	Styrker og svakheter ved datasettet .....	12
4	Teoretisk rammeverk og metode .....	12
4.1	Gjennomgang av teori og metode .....	12
5	Regresjonsanalyse .....	15
5.1	Lønnsnivået er likt i Italia og Norge (A).....	16
5.2	Forskjell i lønnsnivået i Italia og Norge (B).....	18
5.3	Ingen forskjell i avkastning på utdanning for kjønn i Italia og Norge (C).....	19
5.4	Ingen forskjell i avkastning på ferdigheter for kjønn i Italia og Norge (D) .....	20
5.5	Forskjell i lønnsgap mellom offentlig og privat sektor for kjønn i Italia og Norge (E)	21
6	Konklusjon .....	23
7	Litteraturliste .....	24
8	Vedlegg .....	26
8.1	Valg av funksjonsformer .....	26
8.2	Test for heteroskedastisitet .....	27
8.3	Resultater fra regresjonsanalysene .....	28

# 1 Innledning

Temaet for bacheloroppgaven er å analysere lønnsforskjeller mellom kvinner og menn i Italia og Norge. I takt med utviklingen mot et mer likestilt samfunn har stadig flere kvinner satset på egne yrkeskarrierer. I Norge er omtrent like mange kvinner som menn sysselsatte.<sup>1</sup> (SSB; 2018) Til tross for at flere kvinner enn menn fullfører høyere utdanning, og har fått et solid fotfeste i arbeidslivet, tjener kvinner fortsatt mindre enn menn.<sup>2</sup> (NOU 2015:6 (2008, kap 1)) Lønnsgapet mellom kvinner og menn har vært tema for en rekke studier, både nasjonalt og internasjonalt. Likestilling mellom kvinner og menn er et sentralt tema på den politiske agendaen både i Italia og Norge. Lønnsutjevning er videre ansett som et av de viktigste verktøyene i likestillingsarbeidet, og har stor betydning for arbeidstakerens motivasjon. (Lazear & Gibbs, 2009)

Norge har siden tidlig på 1970-tallet jobbet aktivt for likestilling i arbeidsmarkedet. I 1978 vedtok Stortinget Likestillingsloven som skulle fremme likestilling og motarbeide diskriminering av kvinner. I den norske regjeringens politiske plattform legges det til grunn at det skal føres en arbeidslivspolitikkk hvor likestilling og likelønn, utviklingsmuligheter og adgang til kompetansebygging er sentrale elementer.<sup>3</sup> Likestillings- og diskrimineringsloven §34 pålegger arbeidsgiver å lønne kvinner og menn likt uavhengig av kjønn. Loven sier at «kvinner og menn i samme virksomhet skal ha lik lønn for samme arbeid eller arbeid av lik verdi. Lønningen skal fastsettes på samme måte, uten hensyn til kjønn».<sup>4</sup> Til tross for dette, tyder lønnsgapet på at det fortsatt eksisterer en kjønnsbasert lønnsdiskriminering i det norske arbeidsmarkedet.

## 1.1 Problemstilling

Bacheloroppgaven tar for seg lønnsforskjeller mellom kvinner og menn i Italia og Norge, og undersøker hvilke faktorer som påvirker lønnsforskjellene. Landene er valgt på bakgrunn av at de er to europeiske høyinntektsland, men med relativt ulike velferdsordninger.

---

<sup>1</sup> [https://www.ssb.no/befolkning/artikler-og-publikasjoner/\\_attachment/341883](https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/341883)

<sup>2</sup> <https://www.regjeringen.no/contentassets/7234dd2587f94f1d926c79e6b057831f/no/pdfs/nou20152015006000dddpdfs.pdf>

<sup>3</sup> <https://www.regjeringen.no/no/dokumenter/nou-2008-6/id501088/sec2>

<sup>4</sup> [https://lovdata.no/dokument/NL/lov/2017-06-16-51/KAPITTEL\\_2#§8](https://lovdata.no/dokument/NL/lov/2017-06-16-51/KAPITTEL_2#§8)

Bacheloroppgaven har som formål å besvare følgende problemstilling: «Eksisterer det et lønnsgap mellom kvinner og menn i Italia og Norge, og i så fall hvor stort er det?» I tillegg ønsker vi å undersøke hvordan lønnsavkastningen kan forklares av ulike faktorer, og vi har dermed formulert følgende delproblemstilling: «I hvilken grad kan størrelsen på lønnsgapet forklares av utdanningsnivå, arbeidserfaring, kognitive ferdigheter, og valg av sektor?»

Vi skal med andre ord studere i hvilken grad nivået på formell utdanning, antall år med arbeidserfaring, samt observerte kognitive ferdigheter har effekt på timelønnen til kvinner og menn. Med kognitive ferdigheter menes ferdigheter i leseforståelse og tallforståelse. Videre skal vi undersøke om lønnsgapet mellom kvinner og menn varierer i offentlig og privat sektor.

Å sammenlikne lønnsforskjellene mellom kvinner og menn i Italia og Norge er interessant ettersom kvinners deltakelse i arbeidsmarkedet varierer mellom landene. Tall hentet fra OECD viser til at andelen sysselsatte kvinner i Norge i alderen 15 til 64 år var 75.3% mot 55.9% i Italia i 2017.<sup>5</sup> Videre viser tall fra det samme året at lønnsgapet mellom kvinner og menn er høyere i Norge sammenliknet med Italia. I Norge var lønnsforskjellen på 7.12% mot 5.56% i Italia. (OECD; 2017)

For å kunne besvare problemstillingen skal vi utføre deskriptive og økonometriske analyser av datasettet. Vi vil gjennom en deskriptiv analyse forsøke å kartlegge lønnsforskjeller mellom kjønnene i de to landene med hensyn til de variablene vi antar påvirker deres lønnsavkastning. Dette vil være et viktig grunnlag for videre regresjonsanalyser, og vil hjelpe oss med å forstå eventuelle forskjeller i hvilke variabler som påvirker lønnsnivået. Dataanalyseprogrammet Stata er benyttet i utførelsen av regresjonsanalyser og uthenting av deskriptiv statistikk.

## 1.2 Struktur

Oppgavens første kapittel har hittil presentert bacheloroppgavens tema, problemstilling og gitt en innledning til kvinners arbeidsdeltakelse i de to landene. Kapittel 2 presenterer det teoretiske rammeverket og tidligere studier. I kapittel 3 vil datamaterialet og utvalget presenteres, før vi vil presentere funnene fra den deskriptive analysen. Det fjerde kapittelet viser den økonometriske modellen som anvendes i oppgaven, samt hypoteser om hvilke faktorer som

---

<sup>5</sup> <https://stats.oecd.org/index.aspx?queryid=64197>

antas å ha positiv effekt på lønnsnivået. I kapittel 5 vil de empiriske resultatene fra regresjonsanalysen presenteres. Funnene vil diskuteres og drøftes i lys av teori og empiri, før vi i kapittel 6 oppsummerer hovedfunnene i oppgaven.

## **2 Teoretisk rammeverk og tidligere litteratur**

I dette kapitlet presenteres teori og empiri knyttet til lønnsforskjeller mellom kjønn. Vi vil presentere to teoretiske rammeverk: humankapitalteori og preferanseteori, før vi vil gjennomgå relevant empiri om lønnsforskjeller i arbeidslivet. Humankapitalteorien (G. Becker) og preferanseteorien (C.Hakim) ser individets plassering på arbeidsmarkedet som et resultat av valg, og tar utgangspunkt i at alle starter med like muligheter i det de inntar arbeidslivet.

Humankapitalteorien tar utgangspunkt i at når et individ investerer i utdanning øker det produktiviteten, og videre til økt personlig økonomisk avkastning (lønn). Lønn avspeiler med andre ord produktiviteten til arbeidstakeren. Becker peker på at kvinner og menn velger ulike utdanningsretninger, og dette kan være en forklaring på lønnsforskjellene i arbeidsmarkedet. Kvinner tilbringer i gjennomsnitt mindre tid på arbeidsmarkedet enn menn, ettersom kvinner i stor grad har hovedansvaret for barn og husarbeid. Ifølge Becker går kvinner derfor glipp av generell yrkeserfaring, samtidig som den kunnskapen de allerede har foreldes. Når kvinner returnerer til arbeidslivet er de mindre produktive, og muligheten for opprykk og karriereutvikling er svekket. Dette resulterer i at kvinner får lavere avkastning på utdanningen de tar, og av den grunn velger å utdanne seg til yrker hvor kunnskap forringes mindre over tid.

Preferanseteorien peker på at kvinners og menns orientering mot henholdsvis lønnet arbeid og familiearbeid er forskjellig. Teorien hevder at kvinner i stor grad prioriterer familie fremfor karriere, og dermed velger bort karrierebygging.

Et kjønnsdelt arbeidsmarked kan også forklare ulikheter i lønn. Kjønnssdeling i arbeidsmarkedet vil si at kvinner og menn gjennomgående arbeider i ulike yrker og sektorer, og har ulike stillinger i organisasjonshierarkiet. (NOU 2015:6) Teorien hevder at individets plassering på arbeidsmarkedet ikke kan forklares av individuelle valg alene, men som et resultat av strukturer i arbeidsmarkedet. Et kjønnsdelt arbeidsmarked har negative konsekvenser for lønnsfordelingen mellom kvinner og menn. Lønnsforskjellene har sammenheng med såkalt horisontal og vertikal segregering. Horisontal segregering betyr at

kvinner og menn velger ulike yrker, og fordeler seg ulikt på næring, sektor og bransje. Dette betyr at kvinner velger typiske «kvinneyrker» og menn velger typiske «mannsyrker». Typiske kvinneyrker er i helse- og sosial- og undervisningssektoren, mens det er en overvekt av menn i industrien, og i bygg- og anleggsvirksomhet. (SSB; 2014) Typiske kvinneyrker har flatere struktur og færre nivåer i hierarkiet enn hva mannsyrker har. I praksis betyr horisontal segregering at kvinner er underrepresentert i de godt betalte yrkene på arbeidsmarkedet. (Eurofond, 2013) Vertikal segregering handler om posisjoner i organisasjonshierarkiet, og innebærer at kvinner og menn har ulike posisjoner, selv med samme yrkesbakgrunn. (NOU 2008:6, 2008, s.25)

Det er gjort en rekke studier på lønnsforskjeller mellom kvinner og menn. I 2013 publiserte EU-kommisjonen en rapport om likelønn; *Tackling the pay gap between women and men*.<sup>6</sup> Rapporten definerer likelønn mellom kvinner og menn, drøfter årsaker til lønnsgapet, og gir en statistisk oversikt over lønnsforskjeller mellom kvinner og menn i EU-landene. EU-kommisjonen definerer lønnsgapet som den relative forskjellen i gjennomsnittlig timelønn mellom kvinner og menn for hele økonomien. Statistikken viser at det eksisterer store forskjeller i det gjennomsnittlige lønnsgapet mellom EU-landene, og at lønnsgapet øker med alder, utdanning og ansiennitet. Rapporten peker på at de store lønnsforskjellene mellom landene må sees i relasjon til andre forhold i arbeidsmarkedet. Et stort lønnsgap er vanligvis karakterisert ved et arbeidsmarked som er svært kjønnsdelt eller at en stor andel av kvinnene arbeider deltid.

Italias og Norges velferdsordninger kan forklare noe av forskjellene på arbeidsmarkedene. Den norske velferdsstaten tilbyr mange familierettede ordninger, som foreldrepenger, barnetrygd og barnehageplass til alle. Dette styrker muligheten til å kombinere familie- og yrkesliv, og har bidratt til at den norske arbeidsstyrken består av omtrent like mange kvinner som menn.

Det italienske arbeidsmarkedet er derimot preget av et «gammeldags» kjønnsrollemønster, og kvinners sysselsetting er langt lavere enn i Norge. Italia mangler et velferdssystem som gjør det mulig å kombinere familie- og yrkesliv, med mindre man tilhører den velstående overklassen. Moderne kvinner som ønsker å skape en egen yrkeskarriere må av økonomiske hensyn vente i mange år med å få barn. Det at italienske kvinner i økende grad tar utdanning, og er blitt

---

<sup>6</sup> [https://www.equalsalary.org/wp-content/uploads/2018/10/Tackling\\_Gender\\_Pay\\_Gap\\_European\\_Union\\_Report2014\\_ES\\_P21.pdf](https://www.equalsalary.org/wp-content/uploads/2018/10/Tackling_Gender_Pay_Gap_European_Union_Report2014_ES_P21.pdf)

lønnsarbeidere, er en viktig årsak til at Italia er det landet i Europa som har høyest gjennomsnittsalder for førstegangsfødende. (The Independent) Italienske og norske arbeidstakere har med andre ord ulik tilgang til offentlige velferdsordninger, som påvirker muligheten til å kombinere familie- og yrkesliv.

Til tross for disse velferdsordningene, er det norske arbeidsmarkedet kjønnsdelt. Norske kvinner velger i stor grad tradisjonelle «kvinneyrker» og velger i lange perioder å jobbe i redusert stilling. Nesten fire av ti norske kvinner jobber deltid, mens under to av ti menn jobber deltid.<sup>7</sup> (WEF, 2017) Kvinner og menn fordeler seg ulikt mellom sektorer, og også her er forskjellene større i Norge sammenliknet med Italia. I 2015 var 66.67% sysselsatte i offentlig sektor kvinner, mens det tilsvarende tallet for Italia var 55.9%. Videre viser statistikken at i Italia, hvor kvinners sysselsetting er relativt lavere og arbeidsmarkedet er mindre kjønnsdelt, er lønnsgapet mindre enn i Norge. Dette kan reflektere at det er en lav andel med ufaglærte kvinner i yrkeslivet i Italia.

### **3 Datamaterialet**

I dette kapittelet skal vi presentere datamaterialet som er benyttet i bacheloroppgaven. Videre presenteres funnene fra de deskriptive analysene, og til slutt nevnes noen styrker og svakheter med datamaterialet.

#### **3.1 Om datamaterialet**

Dataene vi har brukt i bacheloroppgaven bygger på informasjon hentet fra den internasjonale undersøkelsen om lese- og tallforståelse (PIACC) i regi av OECD. Datamaterialet ble samlet inn i perioden august 2011 til april 2012, og er gjennomført i tilsammen 21 land. Hensikten med undersøkelsen var å finne virkningene av utdanning gjennom å se på tre sentrale kognitive ferdigheter hos mennesker i alderen 16 til 65 år. Ferdighetene ble omtalt som nøkkelferdigheter, og inkluderer leseforståelse, tallforståelse og problemløsning i IKT-miljø. Deltakerne ble også spurt om andre forhold som lønn, utdanningsnivå, kjønn, alder, etnisitet, arbeidsplass mm. Analysene vi har gjort i denne oppgaven baserer seg på disse resultatene, og er hentet ut fra et

---

<sup>7</sup> [http://www3.weforum.org/docs/WEF\\_GGGR\\_2017.pdf](http://www3.weforum.org/docs/WEF_GGGR_2017.pdf)



representativt utvalg av italienske og norske arbeidstakere. (Bjørkeng 2013)

Vi har valgt timelønn som avhengig variabel, og har kun valgt å beholde de respondentene som har oppgitt sin timelønn i spørreskjemaet. Vi antar at det er tilfeldig hvem som har utelatt informasjonen, og at utvalget vårt derfor er representativt for populasjonen. Vi har i tillegg valgt å utelate individer under 25 år og kun fokusere på den såkalte «prime working age». (OECD) Dette er for å unngå skjevheter i gjennomsnittsverdiene. I utvalget vårt har vi inkludert både deltids- og heltidsansatte, og de som jobber i privat og offentlig sektor. Vi har valgt å utelate de som ifølge spørreskjemaet arbeider i non-profit organisasjoner, for eksempel veldedighet og religiøse organisasjoner.

### 3.2 Deskriptive analyser

De deskriptive analysene har som formål å identifisere forskjeller i lønn mellom kvinner og menn i Italia og Norge, og bidra til å forstå funnene i regresjonsanalysen. Analysene inkluderer den avhengige variabelen *timelønn* samt de uavhengige variablene. Timelønn er oppgitt i amerikanske dollar, og justert for kjøpekraftsparitet som gjør det mulig å sammenligne timelønningen på tvers av landegrensene.

Den uavhengige variabelen *kvinne* er en binær variabel, som tar verdien 1 hvis individet er kvinne og 0 hvis mann. Det samme gjelder de binære uavhengige variablene *Italia* og *offentlig sektor*. De uavhengige variablene *utdanning* og *arbeidserfaring* er mål på antall år fullført utdanning og antall yrkesaktive år.

Den uavhengige variabelen *leseferdigheter* er definert som evnen til å forstå og bruke skriftlig tekst for å kunne delta i samfunnet, oppnå egne mål og utvikle sine ferdigheter og sitt potensial. Variabelen *tallforståelse* brukes om evnen til å bruke, tolke og formidle matematisk informasjon, for å kunne håndtere situasjoner i voksenlivet som stiller krav til matematisk kunnskap. (Bjørkeng, 2013) Etersom ingen av individene i det italienske datasettet har angitt svar for problemløsningen innenfor IKT, velger vi å se bort fra denne variabelen. Både leseferdigheter og tallforståelse måles på en skala fra 0 til 500. (Bjørkeng; 2013)

### 3.2.1 Deskriptiv statistikk for avhengig variabel

	Alle		Kvinner		Menn	
	Norge	Italia	Norge	Italia	Norge	Italia
Gjennomsnittlig timelønn	25.51	15.40	23.43	14.92	27.52	15.83
Minimum	5.89	4.03	5.89	4.03	6.2	4.09
Maksimum	64.50	62.22	64.50	62.1	64.37	62.22
Standardavvik	8.6	8.066	7.067	7.66	9.43	8.39
Antall observasjoner	3023	1848	1485	876	1538	972

Tabell 1: Deskriptiv statistikk for lønn

Tabell 1 gir en oversikt over timelønn for kvinner og menn i Italia og Norge. Etter å ha fjernet respondenter som ikke er aktuelle for vår analyse eller som ikke har avgitt noe svar, sitter vi igjen med 4871 respondenter. Av disse er 3023 nordmenn og 1848 italienere.

Vi ser at den gjennomsnittlige timelønnen i Norge er 25.51 US dollar. Norske kvinner har en timelønn som tilsvarer 85% av menn sin timelønn. Tallene stemmer overens med SSB-tall fra 2012, hvor kvinners lønn utgjorde 86.5% av menns.<sup>8</sup> Gjennomsnittlig timelønn i Italia er betydelig lavere enn i Norge, og tilsvarer kun 60% av norsk timelønn. Ser vi derimot på lønnsgapet mellom kvinner og menn i Italia, er det betydelig lavere enn i Norge. Kvinners timelønn tilsvarer 94 % av menns timelønn i Italia. Standardavviket forteller oss videre at det er større spredning i lønn hos menn i begge land.

I den deskriptive statistikken for de uavhengige variablene vil vi inkluderes de variablene vi mener er relevante for vår analyse, og som vi antar har positiv effekt på timelønnen. Vi har valgt å vise både gjennomsnittlige data for hvert land, samt for kvinner og menn separat.

---

<sup>8</sup> <https://www.ssb.no/arbeid-og-lonn/statistikker/lonnansatt/aar/2013-03-20>

### 3.2.2 Deskriptiv statistikk for uavhengige variabler

	Alle		Kvinner		Menn	
	Norge	Italia	Norge	Italia	Norge	Italia
Antall år utdanning	14.66 (2.449)	11.6 (4.064)	14.69 (2.433)	11.50 (3.857)	14.63 (2.464)	11.68 (4.245)
Arbeidserfaring i år	22.69 (11.605)	19.37 (11.747)	20.06 (10.914)	17.32 (11.498)	23.2 (12.015)	21.3 (11.657)
Leseferdigheter	282.92 (46.611)	253.12 (44.832)	280.98 (45.962)	252.84 (43.89)	284.73 (47.145)	253.43 (45.846)
Tallforståelse	284.21 (54.295)	250.79 (50.866)	276.57 (53.109)	244.42 (49.522)	291.31 (54.433)	257.72 (51.414)
Offentlig sektor (prosent)	41.1%	28.9%	65.4%	58.8%	34.8%	41.2%

Tabell 2: Deskriptiv statistikk for utdanning, arbeidserfaring, leseferdigheter, tallforståelse og offentlig sektor i Italia og Norge

Når Italia og Norge sammenlignes finner vi forskjeller i utdanningsnivå. Norske kvinner er i gjennomsnitt litt høyere utdannet enn menn, mens det er motsatt i Italia. PIACC-rapporten understreker at utdanning har en positiv effekt på ferdighetene. Ettersom gjennomsnittlig utdanningsnivå er høyere i Norge, er det naturlig at også ferdighetsnivået er høyere. Videre observerer vi at menn har lengre arbeidserfaring i begge landene, som blant annet kan forklares av preferanseteorien til Hakim.

I PIACC-rapporten kommer det fram at gjennomsnittet for henholdsvis leseferdigheter og tallforståelse er 273 og 269 poeng i OECD-landene. Ut fra den deskriptive analysen kan vi lese at Italia skårer lavere enn gjennomsnittet i lese- og tallforståelse, og at Norge skårer høyere. Videre observerer vi at menn har et høyere ferdighetsnivå enn kvinner i begge land, men denne forskjellen er ikke signifikant når det korrigeres for utdanning, arbeidserfaring og sektor. I Norge finner vi at 41.1% av arbeidsstyrken er ansatt i offentlig sektor, og av disse er 65% kvinner og 35% menn. Denne fordelingen indikerer at Norge er horisontalt segregert, som kan forklare noe av lønnsgapet mellom norske kvinner og menn. I Italia ser vi ikke den samme horisontale segregeringen ettersom kvinner og menn stort sett jobber i privat sektor. I Italia

jobber kun 28.9% av arbeidsstyrken i offentlig sektor, og av disse er 59% kvinner og 41% menn.

### **3.3 Styrker og svakheter ved datasettet**

Datasettet inneholder mange variabler som kan forklare timelønnen, blant annet variabler som utdanning, kjønn, arbeidserfaring, kognitive ferdigheter og opphav. PIACC-undersøkelsen ble gjennomført for et utvalg av populasjonen, og det vil alltid være en risiko for at resultatet ikke er representativt for hele befolkningen. Det store populasjonsutvalget styrker derimot sannsynligheten for at resultatene er gjeldene for hele befolkningen, og er derfor en styrke ved datasettet. Det vil alltid være noen forhold som ikke kan forklares og som blir fanget opp av støyledet, men det omfattende datasettet gjør at en større del av timelønnen kan forklares av variablene.

En svakhet med datasettet er at det ikke tar hensyn til familieforhold og bosted. Generelt sett er lønnsnivået høyere i byene (Michael P. Todardo og Stephen C. Smith, 2015). I tillegg kan forhold som foreldres inntekts- og utdanningsnivå påvirke hvor mye utdanningen til et individ koster av seg. (Kristin F. Butcher og Anne Case, 1994; Hægeland, 2003) Vi skulle også sett at utdanningsvariabelen var mer omfattende. I oppgaven tar vi utgangspunkt i at marginaleffekten av ett års ekstra utdanning er lik for alle individer, men i realiteten er ikke det nødvendigvis sannheten. Det er grunn til å tro at to ulike utdanninger, med samme tidsforløp, kan gi ulik avkastning. I tillegg kan det tenkes at avkastningen av ett år ekstra utdanning for eksempel er høyere for en masterstudent enn en bachelorstudent.

## **4 Teoretisk rammeverk og metode**

I dette kapittelet vil vi gjennomgå regresjonsmodellen som er benyttet i bacheloroppgaven i tillegg til våre utarbeidede hypoteser. Hypotesene er basert på tidligere studier og empiri.

### **4.1 Gjennomgang av teori og metode**

For å undersøke hvilke faktorer som påvirker lønnsnivået må vi definere en lønnslikning. Lønnslikningen vil inneholde de variablene vi mener er relevante for vår analyse. I første omgang definerer vi en forenklet lønnslikning (1).

$$(1) \text{timelønn} = f(\text{female}, x)$$

I denne forenklete lønnslikningen er timelønnen en funksjon av kjønn. Andre kontrollvariabler blir fanget opp av  $x$ . Det er rimelig å anta at  $x$  vil representere humankapitalvariabler og arbeidsrelaterte variabler som utdanning, ferdigheter, arbeidserfaring, sektor osv. Disse faktorene vil som tidligere antatt ha positiv effekt på timelønnen ved at blant annet produktiviteten til arbeidstakeren øker. For å kunne estimere en lønnslikning med våre data kreves en spesifisering av likning (1). En modell som ofte er brukt av økonomer er Mincers lønnsmodell uttrykt ved likning (2)

$$(2) \ln Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 (X_{i3})^2 + \varepsilon$$

I denne modellen ser vi på logaritmen til timelønn ( $\ln Y$ ) som en funksjon av utdanning ( $X_1$ ) og arbeidserfaring ( $X_2$ ) for et individ  $i$ . Arbeidserfaring er uttrykt både lineært og kvadratisk ( $X_3$ )<sup>2</sup> for å ta høyde for mulig ikke-lineær effekt av arbeidserfaring. Økonomiske sammenhenger mellom variabler vil aldri være deterministiske. Vi inkluderer derfor et additivt stokastisk støyledd som fanger opp andre variabler som forstyrrer forholdet mellom  $Y$  og  $X$ , og som ikke er eksplisitt inkludert i modellen. Koeffisientene i regresjonslikningen indikerer stigningstallet og kvantifiserer effekten av uavhengig variabel på avhengig variabel. For eksempel vil koeffisienten foran utdanning,  $\beta_1$ , beskrive effekten av ett år ekstra utdanning på timelønnen gitt at nivået på de andre kontrollvariablene holdes konstant. Tolkningen av dette parameteret er at det (multiplisert med 100) vil beskrive prosentvis endring i timelønn som følge av en enhets endring i utdanning, altså ett år. Vi antar at alle år med utdanning gir lik avkastning. Konstantleddet  $\alpha$ , er det punktet hvor regresjonslinjen skjærer  $y$ -aksen. Det siste leddet i regresjonslikningen,  $\varepsilon$ , er støyleddet.

Den logaritmiske-lineære funksjonsformen er valgt på bakgrunn av at den gir best føyning til datasettet vårt (se vedlegg 8.1). Log-lineære modeller er mye brukt i litteraturen, og tidligere studier viser til at funksjonsformen ofte gir bedre føyning. Videre muliggjør log-lineære formuleringer å sammenlikne estimerte lønnsgap på tvers av land siden estimert lønnsforskjell kan tolkes som tilnærmet prosentvist lønnsgap.

Populasjonens regresjonslinje ( $Y = \alpha + \beta X + \varepsilon$ ) er en ukjent linje som vi ønsker å finne et best mulig estimat for. Ved hjelp av datasettet estimerer og tallfester vi parameterne  $\alpha$  og  $\beta$ , og

estimerer en predikert regresjonslikning for utvalget vårt ( $\hat{Y} = a + bX$ ). På grunn av utvalgsvariansen vil ikke regresjonslinjen til utvalget nødvendigvis sammenfalle med regresjonslinjen til populasjonen. I regresjonslinjen til populasjonen har avhengig variabel en lineær sammenheng med uavhengige variabler, og vi finner det troverdig at antakelsene som må ligge til grunn for bruk av minste kvadraters metode (MKM) er oppfylt. Teorien bak MKM er å minimere summen av det kvadrerte avviket mellom den faktiske verdien,  $Y$ , og den predikerte (observerte) verdien,  $\hat{Y}$ . Ved å minimere summen av det kvadrerte avviket finner vi utvalgsparameterne  $a$  og  $b$ .

$$\min \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2$$

hvor:

$Y$  er den faktiske verdien på  $Y$  for observasjon  $i$

$\hat{Y}$  er den predikerte verdien på  $Y$  for observasjon  $i$ , basert på estimatene for  $\alpha$  og  $\beta$

Avviket mellom faktisk timelønn og predikert timelønn for observasjon  $i$  blir kalt for residualen. Dette er altså den delen av lønnen som ikke blir forklart av vår modell. Forklaringsfaktoren avhenger av hvilken funksjonsform vi tar i bruk.

MKM bestemmer koeffisientene slik at summen av kvadrerte residualer er minst mulig og at støyleddet i gjennomsnitt er lik null. Estimatene på koeffisientene vil vise om det er rimelig å konkludere om kausalitet mellom avhengig og uavhengig variabel. Vi vil da utføre hypotesetester om koeffisientene med et 5% signifikansnivå for å avgjøre om resultatene er statistisk signifikant. For å bruke MKM legger vi noen antakelser til grunn.

For det første antar vi at forklaringsvariablene ikke er stokastiske. Altså at verdiene for  $X_{i1}$  og  $X_{i2}$  er de samme i hvert utvalg. Et gitt individ  $i$ 's lønn og erfaring vil være den samme i hvert utvalg den blir trukket. Variansen for variablene vil dermed være konstant for den underliggende populasjonen. Videre antar vi at det ikke er noen korrelasjon mellom forklaringsvariablene (multikolaritet). I modellen antar vi et stokastisk støyledd og gitte verdier for forklaringsvariabler slik at kovariansen mellom disse må være lik null. Videre antar vi at det ikke er noen korrelasjon mellom støyleddene (autokorrelasjon), at støyleddene er normalfordelte og null i gjennomsnitt. Til slutt antar vi at støyleddene har konstant varians,

altså at det ikke er heteroskedasitet (Jeffrey W. Woolridge, 2016). Heteroskedasitet er et vanlig problem i tverrsnittsdata, og vi vil derfor bruke Mincers lønnspredikasjonsmodell i Stata og deretter teste for heteroskedasitet, slik at vi senere kan korrigere for det. Funnene illustreres i figur 1 i vedlegg 8.2.

## 5 Regresjonsanalyse

Hensikten med denne oppgaven er å studere lønnsgapet mellom kjønn i Italia og Norge. I den deskriptive analysen observerte vi et lønnsgap mellom kvinner og menn i begge land, og at lønnsgapet var størst på det norske arbeidsmarkedet. Vi ønsker videre å undersøke hvilke faktorer som påvirker timelønnen ved hjelp av regresjonsanalyser, og undersøke om det er forskjeller i hvilke faktorer som påvirker timelønnen i de utvalgte landene. Vi vil undersøke om kvinner og menn har ulik avkastning på utdanning og ferdigheter. Til slutt skal vi undersøke om det eksisterer et lønnsgap mellom privat og offentlig sektor, og se hvilken effekt det har på lønnsnivået til kvinner og menn. Basert på tidligere studier og argumentasjon ovenfor benytter vi oss av Mincers-modell uttrykt ved likning (2) tidligere. Vi skal nå estimere 5 varianter med utgangspunkt i likning (2).

$$(A) \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \varepsilon$$

$$(B) \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \delta_1 italia + \rho_0(fem * italia) + \varepsilon$$

$$(C) \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \delta_1 italia + \rho_0(fem * italia) + \rho_1(edu * italia) + \rho_2(edu * fem) + \rho_3(fem * edu * italia) + \varepsilon$$

$$(D) \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \delta_1 italia + \rho_0(fem * italia) + \rho_4(fem * num) + \rho_5(fem * lit) + \rho_6(ita * num) + \rho_7(ita * lit) + \rho_8(fem * ita * num) + \rho_9(fem * ita * lit) + \varepsilon$$

$$(E) \lnwage = \alpha + \delta_0 fem + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \beta_4 num + \beta_5 lit + \delta_1 italia + \delta_2 pub + \rho_0(fem * italia) + \gamma_0(pub * fem) + \gamma_1(pub * italia) + \gamma_2(pub * fem * italia) + \varepsilon$$

Vi har foretatt t-tester som undersøker om estimert lønnsgap mellom kvinner og menn er statistisk signifikant. Nullhypotesen sier at timelønnen for kvinner er lik timelønnen for menn. Dersom t-verdien overstiger kritisk verdi forkastes nullhypotesen, og vi kan konkludere med at det eksisterer et lønnsgap mellom kjønnene.

Vi ser at det estimerte lønnsgapet mellom kvinner og menn i modellene (A) til (E) varierer mellom 12.8 og 18.2%, og er statistisk signifikant i alle modeller. Lønnsgapet er lavest i modell (D) når vi inkluderer kontroll for kognitive ferdigheter, og størst i modell (C) når vi tester for forskjeller i avkastning på utdanning. I analysene B til E, ser vi at lønnsgapet mellom Norge og Italia varierer mellom 44 og 51%. Lønnsgapet mellom de to landene er størst i modell (E) når vi inkluderer kontroll for sektor. Fullstendig resultat ligger vedlagt i tabell 4.

## 5.1 Lønnsnivået er likt i Italia og Norge (A)

I denne modellen tar vi utgangspunkt i at timelønnen er lik for norske og italienske arbeidstakere. Vi skal bruke modellen til å se om det eksisterer et lønnsgap mellom kvinner og menn, vist som prosentvis forskjell i gjennomsnittslønn. I tillegg bruker vi modellen til å tolke den prosentvise endringen i timelønn som følge av ett år ekstra med utdanning og arbeidserfaring. De kognitive ferdighetene gir den prosentvise lønnsøkningen av ett standardavviks økning i testverdien.

Vi starter med å undersøke  $\delta_0$  for å se om det eksisterer et lønnsgap mellom kvinner og menn. Vi etablerer en nullhypotese som sier at timelønnen til kvinner og menn er lik.  $H_0: \delta_0 = 0$ . Alternativhypotesen blir at kvinner har et lavere lønnsnivå enn menn.  $H_A: \delta_0 < 0$ . Vi setter opp en testobservator (TS) for å avgjøre om vi kan beholde nullhypotesen.

$$TS: \frac{b_0 - \beta}{s_b} \sim t(n - k) = \frac{-0.128 - 0}{0.0108} \sim t_{0.05, 4860} = -11.76 < -1.645$$

Vi ser at testobservatoren er langt lavere enn kritisk verdi ved 5% signifikansnivå (hentet fra tabell A1), og vi må forkaste nullhypotesen. Vi kan derfor konkludere med at kvinner har en lavere timelønn enn menn i begge land. Den predikerte timelønnen til kvinner er 12.8% lavere enn for menn, som sammenfaller med funnene fra kapittel 3. Videre skal vi se på koeffisienten



$\beta_1$  som indikerer den prosentvise endringen i timelønn som følge av ett års ekstra utdanning. Ut fra tabellen ser vi at ett år ekstra utdanning gir 7.5% høyere timelønn, og at resultatet er signifikant ved 5% signifikansnivå. Vi kan derfor konkludere med at det eksisterer lønnsinsentiver for utdanning.

Videre observerer vi at ett år ekstra med arbeidserfaring, representert ved  $\beta_2$  har en statistisk positiv effekt på timelønnen med 2,5%. Grunnen til at vi har inkludert arbeidserfaring på kvadrert form er for å teste om arbeidserfaring har en lineær effekt. Vi oppretter en hypotesetest der  $H_0: \beta_3 = 0$  og  $H_A: \beta_3 < 0$  for å se om arbeidserfaring har en lineær eller kvadratisk effekt på timelønnen. Til 5% signifikansnivå godtar vi alternativhypotesen om at arbeidserfaring har en avtagende effekt på timelønnen. Dette kan forklares med at individer blir mindre produktive med årene.

Ferdighetene for tall- og leseforståelse indikerer at timelønnen øker med 1.7 og 2.1% dersom verdien på ferdighetene øker med et standardavvik. Alene kan vi ikke konkludere med at ferdighetene har noe signifikant effekt på timelønnen. Vi vil derfor ved hjelp av en testobservator avgjøre om ferdighetene samlet sett har en positiv og signifikant effekt på timelønnen. Testobservatoren inkluderer en modell uten restriksjoner (uten ferdighetene) og en modell med restriksjoner (med ferdighetene). Testobservatoren blir da:

$$TS = \frac{(R_U^2 - R_R^2)/h}{(1 - R_U^2)/(n - k)} \sim F_{h, n-k} \rightarrow \frac{(0.3604 - 0.365)/2}{(1 - 0.3604)/(3867 - 7)} = 16.58 > 2.0096$$

Testobservatoren overstiger kritisk verdi i t-fordelingen ved 5% signifikansnivå, og vi konkluderer derfor med at de kognitive ferdighetene har en signifikant effekt på timelønnen. Dette kan også gjøres ved hjelp av en F-test, som vi vil benytte oss av senere.

Ved hjelp av modellen finner vi den predikerte timelønnen for kvinner og menn gitt at verdiene for utdanning, arbeidserfaring og kognitive ferdigheter er lik null.

$$\begin{aligned} \text{For kvinner: } \lnwage &= (\alpha + \delta_0) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4 \\ \lnwage &= (1.62 - 0.128) + 0.075X_1 + 0.025X_2 + (-0.0003)X_2 + 0.017X_3 + 0.021X_4 \end{aligned}$$

$$\begin{aligned} \text{For menn: } \lnwage &= \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4 \\ \lnwage &= 1.62 + 0.075X_1 + 0.025X_2 + (-0.0003)X_2 + 0.017X_3 + 0.021X_4 \end{aligned}$$

Vi finner den predikerte timelønnen for kvinner ved å opphøye den naturlige logaritmen  $e$  i 1,492 (1.62-0.128) og for menn  $e^{1.62}$ . Den predikert timelønnen for kvinner vil da være 4.45 amerikanske dollar og 5.1 for menn. I vår analyse, hvor timelønnen er gitt i amerikanske dollar og justert for kjøpekraftsparitet, føler vi det vil være mer hensiktsmessig å studere koeffisientene enn å se på den predikerte timelønnen.

## 5.2 Forskjell i lønnsnivået i Italia og Norge (B)

I denne modellen skal vi undersøke om lønnsgapet kan forklares av at arbeidstakeren er italiensk, og om lønnsgapet mellom kvinner og menn er lavere i Italia. I kapittel 3 observerte vi et lønnsgap mellom Italia og Norge, og at lønnsgapet mellom kjønnene var lavere i Italia. Vi utvider modellen med en binær variabel som indikerer at individet er italiensk, og en interaksjonsvariabel som indikerer at individet både er italiensk og kvinne. Først tester vi hvorvidt om individet er italiensk har signifikant effekt på lønnen. Ved å se på  $\delta_1$  undersøker vi om det faktisk eksisterer et lønnsgap mellom landene. Vi etablerer en nullhypotese hvor lønnsnivået i landene er likt  $\delta_1 = 0$  mot en alternativhypotese hvor  $\delta_1 < 0$ . Nullhypotesen kan forkastes til 1% signifikansnivå, og vi kan konkludere med at det å være italiensk har negativ effekt på timelønnen. I Italia er timelønnen 44.8 % lavere sammenlignet med Norge. Dette stemmer med funnene i kapittel 3.

Koeffisienten  $\rho_0$  forteller at det prosentvise lønnsgapet mellom kvinner og menn er 3.8 % lavere i Italia sammenlignet med Norge. Dette resultatet er kun signifikant ved 10% signifikansnivå. Ettersom vi opererer med et 5% signifikansnivå, kan vi ikke konkludere med at lønnsgapet mellom kvinner og menn er statistisk forskjellig i Italia og Norge. Lønnsforskjellene mellom kvinner og menn blir omtrent som i modell (A) som tyder på at lønnsgapet i liten grad kan forklares av at individet er italiensk.

Lønnslikningen for hvert av individene i modellen:

Norge:

$$\text{For menn} \quad \ln wage = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4$$

$$\text{For kvinner} \quad \ln wage = (\alpha + \delta_0) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4$$

## Italia

For menn  $\ln wage = (\alpha + \delta_1) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4$

For kvinner  $\ln wage = (\alpha + \delta_0 + \delta_1) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_2)^2 + \beta_4 X_3 + \beta_5 X_4$

### **5.3 Ingen forskjell i avkastning på utdanning for kjønn i Italia og Norge (C)**

I denne modellen skal vi undersøke om lønnsgapet kan forklares av systematisk forskjell i utdanningsnivå mellom kjønnene. I den deskriptive analysen observerte vi at den italienske befolkningen i gjennomsnitt er lavere utdannet enn den norske, og at italienske kvinner i snitt er lavere utdannet enn menn. Vi skal i denne modellen undersøke om avkastningen av ett års ekstra utdanning er forskjellig for kvinner og menn, og om dette kan forklare noe av lønnsgapet mellom kjønnene.

Først vil vi undersøke om utdanningsnivå har signifikant effekt på lønnen. Dette gjøres ved å teste  $H_0: \beta_1 = 0$  mot  $H_A: \beta_1 \neq 0$ . ved hjelp av en t-test. Testobservatoren overstiger kritisk verdi og vi forkaster dermed nullhypotesen om at utdanningsnivå er uten betydning for lønnen. Avkastningen på utdanning i modell (C) er tilnærmet  $100 \times 0,0427\%$ . Det vil si at ett års ekstra utdanning gir ca. 4.3% høyere lønn. Videre skal vi se om marginaleffekten av utdanning er forskjellig for kvinner og menn i Norge. Koeffisienten  $\rho_2$  forteller at avkastningen av ett år ekstra utdanning er 0.4% høyere for kvinner. Videre viser koeffisienten  $\rho_3$  at avkastningen av ett år ekstra utdanning er 0.3% høyere for italienske kvinner enn norske. Vi har ikke grunnlag til å si at disse forskjellene er signifikante. Effekten på timelønnen av ett års ekstra utdanning er lavere i Italia ( $\rho_1$ ), men heller ikke i dette tilfellet kan vi ikke konkludere med at forskjellen er signifikant da p-verdien er høyere enn 5%. Ettersom det ikke er noen forskjell i utdanningskoeffisienten for de fire individene vil lønnslikningen være lik som i modell (B).

Estimert lønnsforskjell mellom kvinner og menn har økt til 18,2% i modell C. Med andre ord ser ikke lønnsgapet ut til å kunne forklares av systematiske forskjeller i formelt utdanningsnivå for kvinner og menn. I den deskriptive analysen observerte vi også at gjennomsnittlig utdanningsnivå var høyere for kvinner enn menn i Norge.

## 5.4 Ingen forskjell i avkastning på ferdigheter for kjønn i Italia og Norge (D)

I denne modellen skal vi undersøke om lønnsgapet kan forklares av systematiske forskjeller i ferdigheter. I OECD-rapporten kom det fram at kvinner i gjennomsnitt skårer dårligere på leseferdigheter og tallforståelse enn menn, og at Italia skårer dårligere enn Norge på begge ferdighetene. Vi vil derfor undersøke om avkastningen av ferdigheter er lavere for kvinner sammenlignet med menn, og om avkastningen i Italia er lavere enn i Norge. Vi bruker de standardiserte testskårene for å lettere kunne sammenligne koeffisientene på tvers av landegrensene. Tolkningen av koeffisientene vil da være den prosentvise lønnsøkningen av ett standardavviks økning i testverdien.

De kognitive ferdighetene ser ut til å ha en positiv effekt på timelønnen med henholdsvis 8.1% for tallferdigheter og 0.5% for leseferdigheter. Vi undersøker om ferdighetenes effekt er statistisk signifikant ved å teste  $H_0 : \beta_4 = \beta_5 = 0$  i modell (D) mot alternativhypotesen at minst av en av parameterne  $\beta_3, \beta_4$  er ulik 0. Undersøker ved en F-test og ser at testobservatoren overstiger kritisk verdi i F-fordelingen, se tabell 5 vedlegg 8.3 Vi forkaster dermed nullhypotesen om at de kognitive ferdighetene er uten betydning for lønnen.

Basert på de estimerte parameterne  $\rho_4, \rho_5, \rho_6, \rho_7, \rho_8, \rho_9$  kan vi hverken konkludere med at marginalavkastningen på de kognitive ferdighetene er høyere for menn enn kvinner, eller at avkastningen er høyere i Norge sammenliknet med Italia. Vi konkluderer derfor med at avkastningen av ferdighetene er like for menn og kvinner uavhengig om individet er fra Norge eller Italia. Disse kan heller ikke forkastes når vi ved en F-test ser om variablene samlet har en signifikant effekt. Lønnslikningene vil også her være like som i modell (B) for alle individene. Ser videre at estimert lønnsforskjell mellom kvinner og menn blir 12,3% som er omtrent som i modell (A) og (B). Lønnsgapet mellom kvinner og menn kan dermed i liten grad forklares av systematisk forskjellig nivå på ferdigheter til tross for at ferdigheter i seg selv ser ut til å ha statistisk signifikant effekt på lønnen.

## 5.5 Forskjell i lønnsgap mellom offentlig og privat sektor for kjønn i Italia og Norge (E)

I denne delen ønsker vi å se om jobbsektor har betydning for lønnen, og om lønnsgapet mellom kvinner og menn er forskjellig i de to sektorene. Til slutt skal vi studere lønnsgapet mellom privat og offentlig sektor i Italia, og se om størrelsen på lønnsgapet er ulikt fra Norge.

Vi utvider grunnmodellen med en binær variabel for offentlig sektor, som har verdien 1 hvis individet er sysselsatt i offentlig sektor, og 0 hvis individet arbeider i privat sektor. Videre oppretter vi interaksjonsvariabler for kvinner, Italia og italienske kvinner. Vi kan da se på hvordan lønnsgapet mellom kvinner og menn er i offentlig sektor i Norge, hvordan lønnsgapet mellom offentlig og privat sektor er i Italia, og til slutt om lønnsforskjellen mellom kvinner og menn i offentlig sektor er ulikt mellom landene.

Alfakoeffisienten gir predikert timelønn til en norsk mann i privat sektor. Lønnsgapet mellom offentlig og privat sektor i Norge representeres ved  $\delta_2$ , mens  $\gamma_1$  indikerer om lønnsgapet mellom offentlig og privat sektor er det samme i Italia. For å se om lønnsgapet er ulikt for kjønnene mellom offentlig og privat sektor i Norge ser vi på  $\gamma_0$ , og  $\gamma_2$  for å se om lønnsforskjellen mellom kjønn i offentlig sektor i Italia avviker fra Norge.

Estimert  $\delta_2$  er -0.86 som innebærer at timelønnen i offentlig sektor er 8.6% lavere enn i privat sektor i Norge. Vi etablerer en hypotesetest der nullhypotesen er  $\delta_2 = 0$  mot alternativhypotesen  $\delta_2 < 0$ . Nullhypotesen forkastes til 1% signifikansnivå og vi konkluderer med at individer i offentlig sektor tjener mindre sammenlignet med de som arbeider i privat sektor.

Nå som vi har registrert en lønnsforskjell mellom offentlig og privat sektor undersøker vi om lønnsgapet mellom kjønnene er ulikt i sektorene. Koeffisienten  $\gamma_0$  indikerer at lønnsgapet mellom kjønnene er 6.7% lavere i offentlig sektor. Vi oppretter nullhypotesen om at lønnsgapet er likt i begge sektorer, mot alternativhypotesen om at lønnsgapet er mindre i offentlig sektor. Til 5% signifikansnivå konkluderer vi med at lønnsgapet mellom kvinner og menn er 6.7% lavere i offentlig sektor sammenlignet med privat sektor i Norge.

Når vi analyserer hvordan lønnsgapet mellom offentlig og privat sektor er i Italia studerer vi koeffisienten  $\gamma_1$ . Den forteller oss at individer ansatt i offentlig sektor i Italia tjener 13.5% mer enn ansatte i privat sektor. Dette er differansen mellom koeffisientene  $\delta_2$  og  $\gamma_1$ . Ved en hypotesetest konkluderer vi med at individer i Italia ansatt i offentlig sektor har høyere timelønn enn ansatte i privat sektor.

Til slutt ønsker vi å undersøke om lønnsgapet mellom kjønnene er større eller mindre i offentlig sektor i Italia sammenlignet med Norge. Koeffisienten  $\gamma_2$  indikerer at lønnsgapet mellom kvinner og menn er 2.2% større i Italia. Til 5% signifikansnivå kan vi derimot ikke konkludere med at lønnsgapet mellom kvinner og menn i offentlig sektor er større i Italia. I Italia vil dermed lønnsgapet mellom kvinner og menn i offentlig sektor være 7.4%.

## 6 Konklusjon

Resultatene fra de estimerte lønnsmodellene bekrefter et lønnsgap mellom kvinner og menn både i Italia og Norge, og at denne forskjellen er mellom 12 og 18% avhengig av hvilke variabler vi kontrollerer for. Resultatene viser at lønnsgapet i liten grad kan forklares av systematiske ulikheter mellom kjønnene i arbeidserfaring, formelt utdanningsnivå eller målte kognitive ferdigheter.

Resultatene viser videre at noe av lønnsgapet på det norske arbeidsmarkedet skyldes at kvinner i større grad enn menn er sysselsatt i offentlig sektor. I den deskriptive analysen observerte vi at 65.4% av sysselsatte i offentlig sektor er kvinner, og regresjonsanalysen estimerte at å jobbe i offentlig sektor har en signifikant negativ effekt på timelønnen med 8.6%. Resultatene fra regresjonsanalysen viser at lønnsgapet mellom kvinner og menn er lavere i offentlig sektor i begge land. Vi kan videre ikke konkludere med at lønnsgapet er lavere i Italia sammenliknet med Norge. På bakgrunn av de resultatene vi har kommet fram til tyder det på at lønnsgapet ikke varier systematisk med formelt utdanningsnivå eller målte kognitive ferdigheter, men at lønnsgapet varier mellom de ulike sektorene.

Bacheloroppgaven tar kun for seg et begrenset utvalg av variabler, noe som gjenspeiles i regresjonsanalysenes forklaringsgrader. I regresjonen for den første modellen er forklaringsgraden ganske lav med en verdi på 36.5%. Denne øker til 54.2% i den siste modellen, som betyr at modellen kan forklare nesten halvparten av variasjonen i timelønnen. For å øke forklaringsgraden ytterligere kunne det vært interessant å inkludere andre relevante variabler, for eksempel bosituasjon, stilling, forpliktelser i hjemmet, partner, antall barn, bransje og yrke.

## 7 Litteraturliste

Begg, D., Vernasca, G., Fischer, S., og Dornbush, R. (2014): *Economics*. 11th edition, McGraw-Hill Education.

Bjørkeng, B. (2013): Ferdigheter i voksenbefolkningen. Resultater fra den internasjonale undersøkelsen om lese- og tallforståelse (PIAAC). Rapporter 42/2013. Statistisk Sentralbyrå.

Bye, Torstein. «Dette er kvinner og menn i Norge» SSB, februar 2018. Hentet 25.03.2019 fra [https://www.ssb.no/befolkning/artikler-og-publikasjoner/\\_attachment/341883](https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/341883)

Bø, Tor Petter: «Høy yrkesdeltakelse blant kvinner i Norden». SSB, mars 2004. Hentet 14.03.2019 fra [https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/hoy-yrkesdeltakelse-blant-kvinner-i-norden?fbclid=IwAR1P7POHlNzjg4whxe9hrfxRD39vDXoPDmw34bH8oQRX\\_dYr\\_0h1cx7WpkQ](https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/hoy-yrkesdeltakelse-blant-kvinner-i-norden?fbclid=IwAR1P7POHlNzjg4whxe9hrfxRD39vDXoPDmw34bH8oQRX_dYr_0h1cx7WpkQ)

Hægeland, Torbjørn. «Økonomisk avkastning av utdanning» i Utdanning 2003. SSB, desember 2003. Hentet 08.03.2019 fra <https://www.ssb.no/a/publikasjoner/pdf/sa60/kap-10.pdf>

Kristoffersen, Sigrun. «Lønnsforskjellene mellom kvinner og menn vedvarer». SSB, desember 2017. Hentet 09.03.2019 fra <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lonnsforskjellene-mellom-kvinner-og-menn-vedvarer--331542%20>

NOU 2008: 6 (2008): *Kjønn og lønn - Fakta, analyser og virkemidler for likelønn*. Oslo: Barne- og likestillingsdepartementet. <https://www.regjeringen.no/contentassets/39b49bdffe6f44faa80f6c3f26de2b8a/no/pdfs/nou200820080006000dddpdfs.pdf>

NOU 2015: 6 (2015) *Grunnlaget for inntektsoppgjørene 2015* Oslo: Departementenes sikkerhets- og serviceorganisasjon. <https://www.regjeringen.no/contentassets/7234dd2587f94f1d926c79e6b057831f/no/pdfs/nou201520150006000dddpdfs.pdf>



OECD. (2017). *Gender wage gap*. Hentet 05.03.2010, fra <https://www.oecd.org/gender/data/genderwagegap.htm>

OECD. (2017). *LFS by sex and age - indicators: Labour force participation rate*. Hentet 7.mars, 2019 fra <https://stats.oecd.org/index.aspx?queryid=64197>

OECD. (2017). *Part-time employment rate*. Hentet 08.03 2019 fra <https://data.oecd.org/emp/part-time-employment-rate.htm>

Rustad Holseter, Anne Marie. «Norge i Europa – statistikk, utdanningssystemer og høyere utdanning» i Utdanning 2002. SSB. Hentet 08.03.2019 fra <https://www.ssb.no/a/publikasjoner/pdf/sa74/kap-12.pdf>

Thomas, R (2005). Using Statistics in Economics. New York: McGraw-Hill Education.

## 8 Vedlegg

### 8.1 Valg av funksjonsformer

Vi sammenligner ulike funksjonsformer for å finne regresjonslikningen som gir best føyning til datamaterialet. Vi gjennomfører en regresjonsanalyse i Stata hvor vi har tar utgangspunkt i variablene timelønn og antall år utdanning,

For å finne den funksjonsformen som best beskriver relasjonen mellom disse to variablene sammenligner vi SSR-verdiene. Det er ikke mulig å sammenligne SSR-en («sums of residuals») i en lineær og logaritmisk modell, fordi de har ulik verdi på avhengig variabel. Dette har vi tatt høyde for ved å benytte Zarembka-metoden. Vi har benyttet den avhengige variabelen (Y) sitt skalerte gjennomsnitt, som gjør at vi kan sammenligne SSR-verdiene. (R L Thomas, 2005) Funksjonsformen med lavest SSR vil være den modellen som gir best føyning til vårt datasett. SSR er et mål på avviket mellom den estimerte modellen og faktiske observasjoner, og er gitt ved  $\sum e_i^2$ . Modellen med lavest SSR vil dermed ha minst variasjon.

Funksjonsform	Formel	SSR
Lineær	$Y = \alpha + \beta X$	1115.929
Lin-log	$Y = \alpha + \beta \ln(X)$	1134.034
Invers modell	$Y = \alpha + \beta \frac{1}{X}$	1177.854
Log-lin	$\ln Y = \alpha + \beta X$	877.365
Log-log	$\ln Y = \alpha + \beta \ln(X)$	883.583

Tabell 3: Funksjonsformene

Vi ser at log-lin modellen gir best føyning til vårt datasett. Dette stemmer med modellen til Jacob Mincer, og vi vil bruke denne for vårt datasett.

## 8.2 Test for heteroskedastisitet

Dersom støyleddene ikke har lik varians, oppstår heteroskedastisitet, og forutsetningene for MKM er ikke oppfylt. Konsekvensen av varierende varians er at estimatoren ved bruk av MKM ikke gir lavest varians, som gjør at estimatoren for variansen i støyleddet og parameterne blir feil. For å teste for heteroskedastisitet benytter vi oss av Breusch-Pagan testen. Testen kontrollerer om forutsetningen om at alle støyledd har konstant varians holder. Vår nullhypotese er at det er homoskedastisitet, altså konstant varians. Alternativhypotesen er heteroskedastisitet, altså ulik varians.

Resultatene fra testen (utført i Stata) blir som følger:

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of lnhour	
chi2(1)	= 40.87
Prob > chi2	= 0.0000

Figur 1: Breusch-Pagan test

Som testen ovenfor viser ble resultatet av testen en p-verdi tilnærmet lik null. Vi må forkaste nullhypotesen om homoskedastisitet, og konkluderer med at vi har heteroskedastisitet. Dette betyr at forklaringsvariablene vil påvirke variansen. For å korrigere for dette, velger vi å benytte oss av robuste standardavvik i regresjonsanalysene våre.

### 8.3 Resultater fra regresjonsanalysene

Tabell 4: Komplette modellestimeringsresultater. Estimerte standardavvik i parentes

	(1)	(2)	(3)	(4)	(5)
	A	B	C	D	E
VARIABLES	lnhour	lnhour	lnhour	lnhour	lnhour
Kvinner	-0.128*** (0.0108)	-0.129*** (0.00997)	-0.182*** (0.0643)	-0.123*** (0.0103)	-0.141*** (0.0154)
Utdanning	0.0751*** (0.00187)	0.0446*** (0.00193)	0.0427*** (0.00317)	0.0450*** (0.00193)	0.0416*** (0.00177)
Erfaring	0.0254*** (0.00200)	0.0231*** (0.00173)	0.0231*** (0.00174)	0.0228*** (0.00174)	0.0215*** (0.00160)
Erfaring kvadrert	-0.000296*** (4.25e-05)	-0.000322*** (3.73e-05)	-0.000323*** (3.74e-05)	-0.000317*** (3.75e-05)	-0.000298*** (3.56e-05)
Tallferdigheter	0.0168 (0.0117)	0.0562*** (0.0101)	0.0566*** (0.0101)	0.0811*** (0.0186)	0.0609*** (0.00963)
Leseferdigheter	0.0208* (0.0116)	0.0174* (0.00984)	0.0174* (0.00985)	0.00558 (0.0185)	0.0156* (0.00934)
Italia		-0.448*** (0.0159)	-0.440*** (0.0639)	-0.439*** (0.0170)	-0.513*** (0.0158)
Italiensk kvinne		0.0380* (0.0196)	0.00697 (0.0196)	0.0268 (0.0196)	0.0196 (0.0196)

	(0.0203)	(0.0911)	(0.0213)	(0.0232)
Kvinne tallf.			-0.0277	
			(0.0241)	
Kvinne lesef.			0.00763	
			(0.0235)	
Italia tallf.			-0.0299	
			(0.0298)	
Italia lesef.			-0.00349	
			(0.0297)	
Italia kvinne tallf.			0.0116	
			(0.0394)	
Italia kvinne lesef.			0.0368	
			(0.0388)	
Utdanning kvinne		0.00357		
		(0.00427)		
Utdanning Italia		-0.00121		
		(0.00475)		
Utdann. Ita. kvinne		0.00314		
		(0.00662)		
Offentlig sektor				-0.0863***
				(0.0178)

Italia offentlig					0.221***
					(0.0298)
Kvinne offentlig					0.0668***
					(0.0240)
Kvinne off. Italia					-0.0218
					(0.0404)
Konstantledd	1.622***	2.267***	2.296***	2.258***	2.354***
	(0.0351)	(0.0354)	(0.0503)	(0.0355)	(0.0328)
Observasjoner	4,867	4,867	4,867	4,867	4,867
R <sup>2</sup>	0.365	0.530	0.531	0.531	0.542

---

*Robuste standardavvik i parentes*  
\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Hypotesetest	Frihetsgrader	t-verdi	p-verdi
(D) $\beta_4 = \beta_5 = 0$	F(2, 4852)	59.54	0.0000
(D) $\rho_4 = \rho_5 = 0$	F(2, 4852)	1.85	0.1567
(D) $\rho_6 = \rho_7 = \rho_8 = \rho_9$	F(4, 4852)	1.61	0.1687

*Tabell 5: F-tester*